

הערות כלליות

1. יש לקרוא בתשומת לב את התרגיל, לענות על כל השאלות התאורטיות ולבצע את כל המטלות המעשיות ככתבם ולשונם.
2. יש להגיש את התרגיל בזוגות אך לא בקבוצות גדולות יותר.
3. המימוש חייב להתבצע בשפת C++. הקומפיילר שבו תתבצע הבדיקה יהיה [Microsoft Visual Studio 2008](#).
4. יש להגיש את התרגיל על גבי דיסק cd-rom לתא התרגילים של הקורס.
5. המימוש של האלגוריתמים חייב להיות כולו שלכם, אין אפשרות להשתמש בפונקציות עזר, מבני נתונים (למעט string) או אלגוריתמים המסופקים על ידי ספריות.
6. ניתן להשתמש בפונקציות ספריה כדי לקרוא ולכתוב קבצים, לבחור מספרים באקראי ולמדוד את זמן הריצה.
7. החומר המוגש על גבי הדיסק חייב להכיל את הקוד, את התשובות לשאלות התאורטיות ואת הגרף המסכם את עילות זמני הריצה של האלגוריתמים השונים.

1. ממשו תכנית המקיימת את הדרישות הבאות:

- 1.1. התכנית מקבלת כקלט קובץ המכיל מילים באנגלית (מילה בכל שורה), את מספר המילים בקובץ ואת גודל המדגם $0 < k \leq n$
- 1.2. התכנית בוחרת באקראי k מילים מתוך הקלט ומדפיסה אותם ל-standard output.

הפקודה להרצת התכנית חייבת להיות:

```
Sample <word file> <number of words> <sample size>
```

לדוגמה, כדי לבחור 100 מילים מתוך הקובץ mywords.txt שמכיל 300 מילים, נרשום:

```
Sample mywords.txt 300 100
```

2. ממשו תכנית המקיימת את הדרישות הבאות:

- 2.1. התכנית מקבלת כקלט קובץ המכיל מילים באנגלית (מילה בכל שורה) ואת מספר המילים בקובץ.
- 2.2. בנוסף התכנית מקבלת כקלט שם של אלגוריתם מיון מתוך השמות הבאים:
insertion, quick, quick-3median, quick-insertion
- 2.3. התכנית ממיינת את המילים בקובץ בסדר לקסיקוגרפי עולה על פי האלגוריתם שנבחר:

insertion - אלגוריתם insertion sort כפי שלמדנו בכיתה
quick - אלגוריתם quick sort כפי שלמדנו בכיתה
quick-3median - אלגוריתם quick sort שבו ה-pivot נבחר על ידי הטכניקה הבאה (median of three):
לוקחים את האיבר הראשון, האמצעי והאחרון, ובחרים את החציון מבין שלושת איברים אלה.
quick-insertion - אלגוריתם זה בנוי משני שלבים:
שלב 1. quick-sort אשר מפסיק למיין כאשר גודל המערך פחות מקבוע k (את ערך הקבוע עליכם לבחור

בעצמכם).
שלב 2. הרצה של אלגוריתם insertion-sort על המערך ה-"כמעט ממוין".

ענו על השאלות הבאות:

- א. הסבירו מדוע יעילות quick-insertion כפונקציה של k היא $O(nk + n \log(n/k))$.
- ב. תארו כיצד תבחרו את k כדי לקבל את זמן הריצה המהיר ביותר.

2.4. התכנית תכתוב את הפלט (המילים הממוינות) לקובץ `sortedwords.txt` ואת הזמן שלקח למיין את המילים (במילי-שניות) ל-`standard output`.

הפקודה להרצת התכנית חייבת להיות:

```
Sort <word file> <number of words> <algorithm name>
```

לדוגמה, כדי למיין את המילים בקובץ `mywords.txt` באמצעות אלגוריתם insertion sort נרשום:

```
Sort mywords.txt 200 insertion
```

3. מלאו את הטבלה הבאה (ציינו את זמני הריצה במילי-שניות):

גודל הקלט	quick-insertion	quick-3median	quick	insertion
4000				
8000				
16000				
32000				
64000				
128000				
256000				
512000				

היעזרו בקובץ [words.txt](#) המכיל 512000 מילים אנגליות.

כדי למלא כל תא בטבלה הריצו את האלגוריתם 3 פעמים, בכל פעם בחרו את הקלט באקראי (באמצעות התכנית שכתבתם בשלב 1) וחשבו את ממוצע הזמנים. לדוגמה, כדי למלא את התא עבור מיון 8000 מילים באמצעות אלגוריתם quick השתמשו בתכנית `Sample` כדי לקבל מדגם בגודל 8000 מילים, מיינו אותו באמצעות `Sort`

ומדדו את הזמן¹. חזרו על פעולה זו עוד פעמיים ורשמו את הזמן הממוצע בטבלה.

אם זמן הריצה עבור גודל מסוים הוא יותר מדקה אין צורך להריץ את האלגוריתם על קלט גדול יותר.

סכמו את המדידות בגרף המתאר את זמני הריצה של כל האלגוריתמים כפונקציה של גודל הקלט.

כדי למדוד את הזמן היעזרו בהסבר של סעיף ז' בשאלה 1 בנושא "זמני ריצה של פונקציות איטרטיביות" בחוברת¹ התרגילים